Robotic Automation for Grasping and Insertion Connectors: A Case Study

Hsien-I Lin, Fauzy Satrio Wibowo, Arthur Peng, and Subhahit Nanda

***Abstract*— Despite advances in robotic automation, the grasping and insertion of connectors remains problematic. This can mainly be attributed to the fact that most connectors have only a few image features by which they can be identified and the wires lack sufficient rigidity to support the connector in a stationary position. Many robots use RGB-D structured light systems to identify the pose of an object and grasp it; however, this type of camera is too expensive for most manufacturing applications. In this study, we developed an inexpensive system in which a pair of two-dimensional industrial cameras provide top and side views to facilitate the grasping and insertion of connectors. The objective functions were to maximize reliability while minimizing cost. In experiments involving 32 trials, the proposed scheme achieved a success rate of 90%.**

Index Terms*— connector, grasping, insertion, connector, 2D camera

I. Introduction

THIS work focuses on the problem of a robot grasping a connector and inserting it into a socket. The grasping of objects by a robot is a challenging problem involving the identification of the connector as well as planning and control of subsequent movements. This issue is particularly difficult when the objects to be grasped are held aloft in the air by wires rather than fixed on the ground.

The most common approach to resolving the problem of object identification involves the use of image processing to estimate the pose of the object. Most pose estimation schemes are based on object detection and feature mapping. Choi *et al.* [1] proposed a 2-point algorithm in which an RGB-D camera is used to extract features via the SIFT algorithm to estimate the physical parameters of the object using a depth map. A perspective n-point algorithm is then used to match the coordinates of a two-dimensional (2D) image frame with the corresponding three-dimensional (3D) frame and compute the pose. Choi *et al.* [2] proposed a 3D model-based approach to extract the features and edges of an object using the SIFT algorithm to derive the physical parameters. Finally, the pose is computed using the perspective n-point algorithm. Song *et al*. [3] extracted the center point of the object and then used a stereo vision technique to calculate the pose.

Structured light cameras use specific patterns of light and 2D cameras to detect depth and obtain the 3D coordinates of objects. They are widely used in machine vision systems, including the Kinect system used in video games and computer technology. A number of high-quality structured light sensors are available from manufacturers such as Gocator Inc. [4] and Solomon Inc. [5]; however, these systems are very expensive. Yang *et al*. [6] designed a stereo-structured light sensor for robotic welding; however, their system was hindered by poor contrast, reflections from metallic surfaces. Chen *et al.* [7] developed an imaging system for random bin-picking systems using a structured light camera for 3D scene acquisition. They found that grasping efficiency could be enhanced by reducing the 3D scene rescanning time.

A common approach to object pose estimation involves the use of structured light cameras to build object point clouds from which to formulate 3D features [8-10]. The viewpoint feature histogram (VFH) [11] and partition viewpoint feature histogram (PVFH) [8] have been applied to random bin-picking tasks by identifying objects via six degrees of freedom (6DoF). Point cloud systems have also been developed using iterative closest point (ICP) registration for pose estimation [12]. The ICP algorithm remains the most commonly-used approach to pose estimation, and researchers have developed a number of variations including Scale ICP [13], HT-ICP [14], and outlier robust ICP [15].

This paper deals with the issue of grasping plain (*i.e.*, featureless) connectors and inserting them onto a designated pin. This task poses three main difficulties: (1) detecting connector features, (2) building point clouds for point-cloud registration to estimate the pose of objects using structured light cameras, and (3) overcoming the need for high-quality structured light cameras. The proposed system uses two cameras (top-view and side-view) which provide constraints aimed at ensuring that the pose of a connector falls within the tolerance of the gripper. After the connector is grasped, robot trajectories and a rotating table are used to control connector insertion.

The remainder of this paper is arranged as follows. Section II presents a problem statement and hardware description. Section III details the methodology. Experiment results and a brief discussion are presented in Section IV. Conclusions are presented in Section V.

II. Problem Statement and Hardware description

The problem addressed in this study was to grasp the connector of a speaker in an arbitrary orientation and insert it onto a desired pin, as shown in Fig. 1. The orientation of the pin to which the connector is attached requires a turntable to rotate the speaker.

Hsien-I Lin is with the National Taipei University of Technology, Taipei 10608 Taiwan (e-mail: sofin@ntut.edu.tw).

Fauzy Satrio Wibowo is with the National Taipei University of Technology, Taipei 10608 Taiwan (e-mail: fauzy.satriowibowo@gmail.com).

Arthur Peng is with the Wistron Corporation, 11469 Taipei, Taiwan (e-mail: arthur_peng@wistron.com).

Subhahit Nanda is with the National Taipei University of Technology, Taipei 10608 Taiwan (e-mail: subhajit27051996@gmail.com).

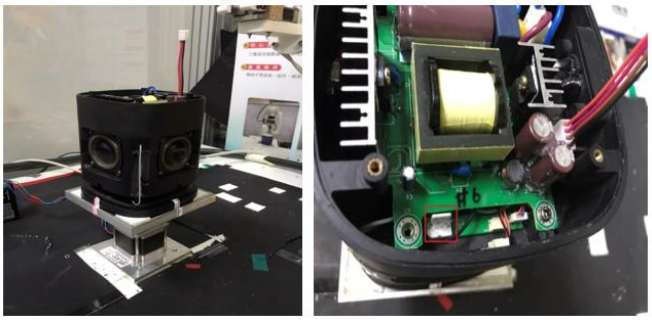


Fig. 1. Primary tasks: (a) grasping of connector; (b) insertion over pin

The hardware used in this work can be divided into four systems: (a) vision system, (b) robotic manipulator, (c) gripper, and (d) rotating table. The vision system employed a full-color industrial camera and lens (Imaging Source DFK 23GP031 GigE). The robotic manipulator was a 6DoF TX40 robot with the CS8C controller (Staubli Inc.). The gripper was a LEHZ 20 Electrical gripper with LECP6 controller. Note that 19N of force was used for grasping. The rotating table employed a NEMA 13 stepper motor and DM542 stepper motor driver controlled using an Arduino Uno. Figure 2 illustrates the hardware installation.

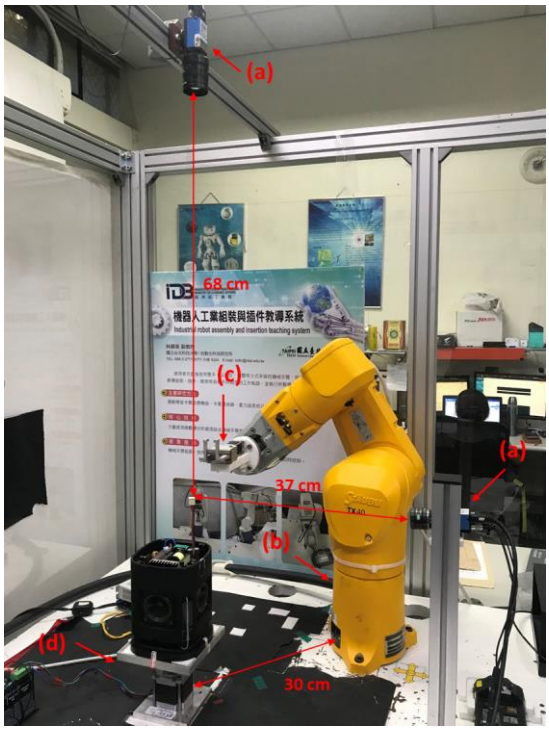


Fig. 2. Hardware installation: (a) vision system, (b) robotic manipulator, (c) gripper, and (d) rotating table

Note that the small size of the connector (5mm x 10.50mm x 12.50mm) increases the difficulty of the process. Figure 3 presents side- and top-views of the connector. Figure 4 presents a top-view of the PCB. According to Fig. 3, there is 4.50mm (A) of space to perform the grasp. Note that a successful grasp requires error of less than 0.8 mm, which poses enormous challenges to the design of the gripper, the vision system, and the manipulation system. Note also that the structure of the connector is averse to automation. A successful cycle requires that the connector be grasped by matching notches (1) and (2) in Fig. 3 with spaces (3) and (4) in Fig. 4. Successful completion of insertion requires pushing of pin (5) in Fig. 4. The design of the gripper fingers is meant to satisfy the connector and gripper constraints. Figure 5 presents a side-view of the gripper finger designed to grasp the connector firmly by matching space (7) in Fig. 5 with connector (10) in Fig. 10. The length of the finger claw was 6 mm because the length of the connector was 5mm.

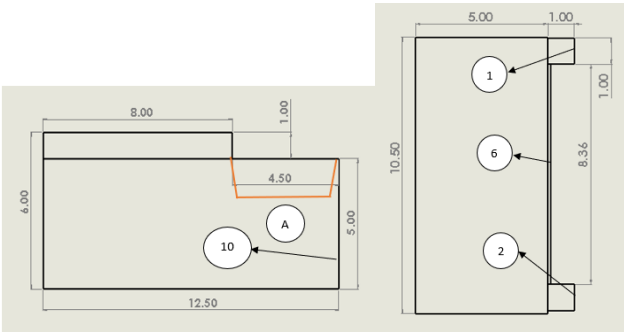


Fig. 3. Side-view (left) and top-view (right) of connector (mm)

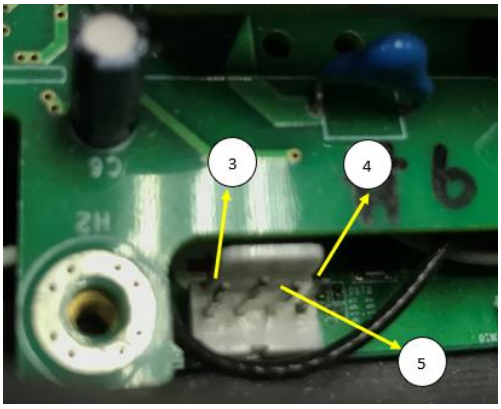


Fig. 4. Top-view of PCB

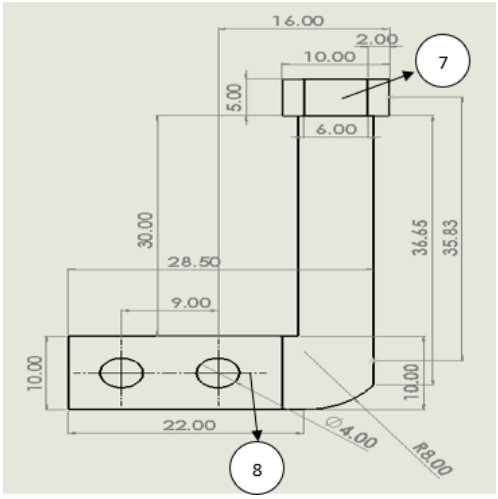


Fig. 5. Side-view of gripper finger (mm)

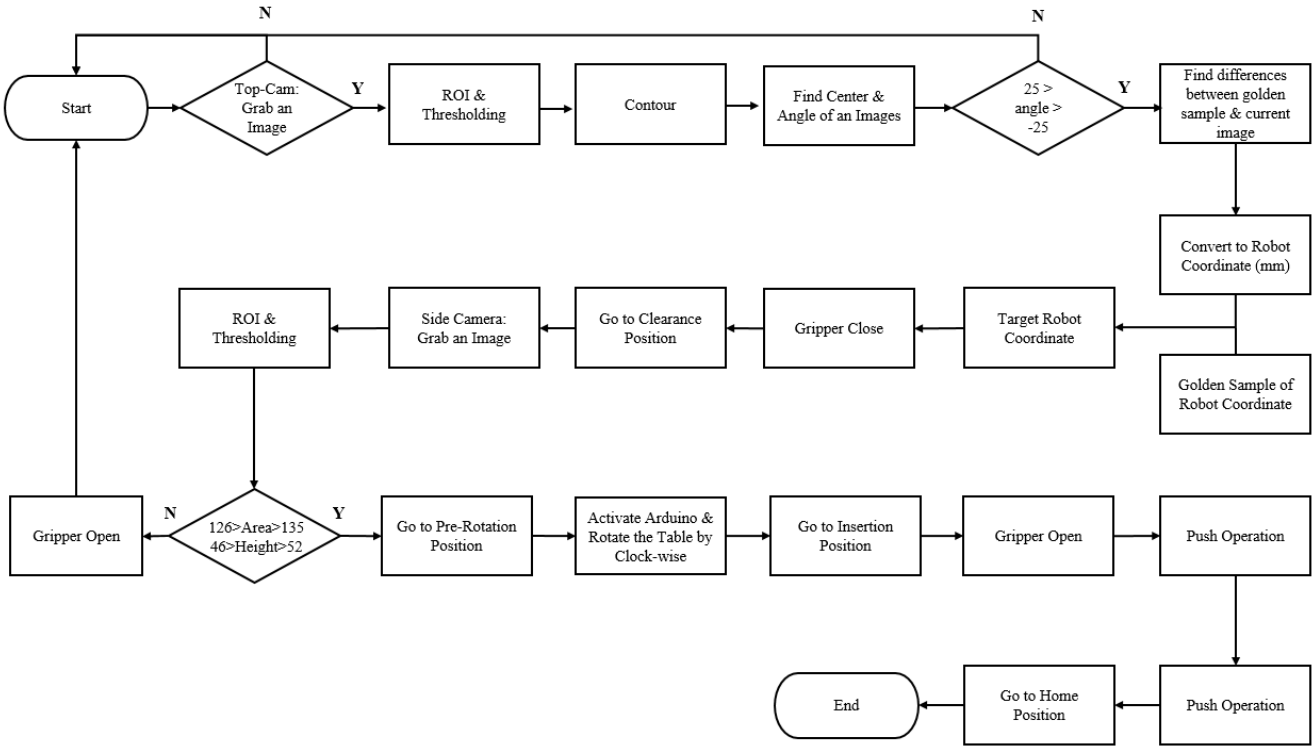


Fig. 6. Flowchart of proposed method

III. METHODOLOGY

Figure 6 presents a flowchart of the methodology used in this study. We first obtained a ground-truth (i.e., ideal) image of the connector in which the orientation is null. The coordinate in the camera pixel coordinate (x_{ground},y_{ground}) and the corresponding robot coordinate are taught. During testing, a current image frame is grabbed and subjected to image processing to calculate the connector coordinate in the camera pixel coordinate ($x_{current}$,$y_{current}$) and the orientation angle (θ). The difference between the ground-truth coordinate and the current coordinate in the image frame is computed and transformed to the robot coordinate.

In the following explanation, we divide the process into (a) the vision system and (b) the robotic manipulation system.

(a) Vision system

An image (1600 x 1200 px) is first captured using a high-quality industrial camera. The captured 64-bit RGB image is converted into a gray-scale image and various filters are used to remove the noise. The gray-scale image is then converted into a binary image using a simple thresholding technique. Note that if no object is present in the scene, then the process is terminated.

In this study, there was no need to work with a full-size image; therefore, we cropped the image to include only the region of interest (ROI), as follows: x = 728 px, y = 528 px, width = 137 px, height = 226 px (X-axis – 9cm, Y-axis – 5cm). Pixels with intensity or grey level within (105, 200) were classified as white and the remainder as black. Figure 7 presents the output of this binary thresholding.

Fig. 7. Binary thresholding

After binary thresholding, the contour area is used to filter out background noise. Within the image frame, the area covered by the contours of the connector varies between 500 and 1150 pixels. From the contours, we calculate the 2D image coordinates ($x_{current}$,$y_{current}$) indicating the center and its rotation angle (θ) (theta). This is performed using the ground-truth image (x_{ground},y_{ground}) and the current image ($x_{current}$,$y_{current}$) to find the center of the object in both image frames. After calculating the angle and center coordinates in the current image frame, we calculate the difference between the two centers in terms of pixels. The output of this process is presented in Figure 8.

The height and rotation angle of the connector along the Y-axis of the connector are calculated from side-view images captured using another camera with resolution of 640 x 480 px, as shown in Fig. 9 (a). Figure 9 (b) shows the ROI of the second camera measuring x = 126 px, y = 105 px, width = 374 px, and

height = 190 px. Using this ROI, we calculate height (H) and grasp quality.

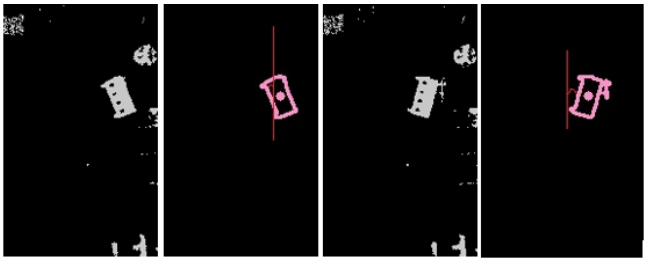
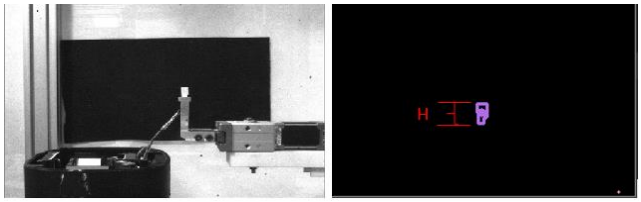


Fig. 8. Center and rotation angle



(a) (b)

Fig. 9. (a) Image captured using side camera; (b) contour

(b) Robot manipulation system

In the following, we present the process flow involved in manipulation of the robot. Before initiating the process, we set a home coordinate to which the robot will return after completing the connection process. Following calculation of the pixel difference, the values are converted from pixels into mm using the following equation:

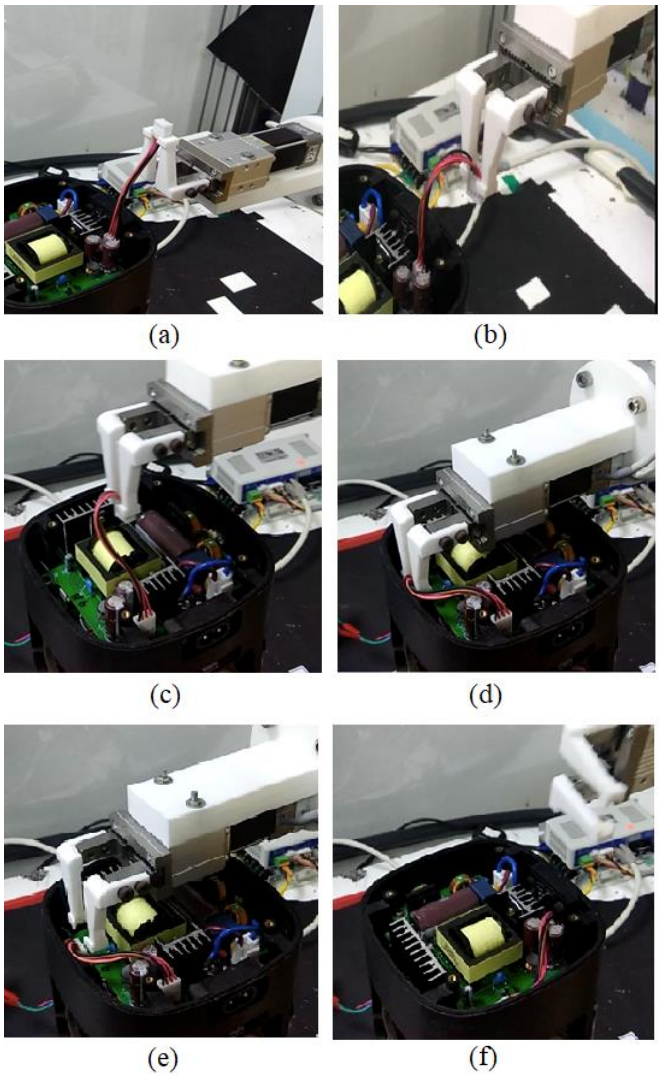
$$\left(x_{difference\ (robot)}\right) = \left(x_{difference}\right) * 0.3127375\ ,$$
$$\left(y_{difference\ (robot)}\right) = \left(y_{difference}\right) * 0.25824\quad .\qquad (1)$$

We add this difference in terms of mm to the ideal coordinates ($\boldsymbol{R_{ideal}}$) to calculate the target coordinates ($\boldsymbol{R_{target}}$). We command the robot to move to the target coordinate to grasp the connector. The robot then follows a pre-defined trajectory to complete the process. It first moves to a position with sufficient clearance to enable rotation of the 6th joint 180 degrees (see Fig. 10 (a)). The robot then moves to a position that allows for rotation of the table (see Fig. 10 (b)). The table is activated via serial communication with the Arduino to rotate the speaker 90 degrees clockwise (see Fig. 10 (c)). The robot then moves the manipulator over the insertion point to insert the connector to the pin (see Fig. 10 (d)). The robot then returns to its home position (see Fig. 10 (e)). After the process has finished, the rotating table rotates 90 degrees anti-clockwise, moving the speaker back to the previous position (see Fig. 10 (f)).

IV. Experiment Results

The proposed scheme was evaluated with the connector starting in various orientations (25> θ >-25) and positions (X, Y) within the ROI. The results of 32 trials are tabulated in Table I. Note that the initial orientation of the connector was represented by random variable p. The time required for the completion of each operation is tabulated in Table II. For applications involving mass production, increasing the length of the wire allows for a corresponding increase in the robot operating speed, reducing the entire process to under 20 seconds (table rotation included). If the rotating table is eliminated, the process can be completed within 15 seconds.

There were a number of constraints on this experiment. The ROI of the grasping system measured just 9 x 5 cm. If the connector were not present in the ROI, then the program would terminate the process and begin the next trial. The proposed scheme is based on 2D tracking and had clearance of only 4.5 mm for grasping; therefore, the system would have been unable to complete the process if the variation in connector height exceeded 2 - 3 mm. In terms of pixels, this means that the connector can be successfully grasped and inserted as long as its height is between 46 and 52 pixels with an area between 126 and 135 (px^2).



(a) (b) (c) (d) (e) (f)

Fig. 10. Grasping and insertion of connector in various steps

Note also that there was a limit on the rotation angle (θ) (theta) of the connector along the Z-axis. Due to limitations on

the working envelope of the robot, the graspable rotation angle of the connector ranges from -25 degrees to 25 degrees around the Z-axis. In the event that the rotation angle of the object in the current frame exceeds this limit, then the program terminates the process and begins the next trial. Due to the length of the connector wire and current configuration of the robot, it operated at only 5% of its maximum speed. Note that even at this speed, the connector wires were under high tension.

We have a number of suggestions to make the assembly process faster, more accurate, and more robust. First, the length of the connector should be increased by 1 - 1.5 cm to provide more flexibility in the movement of the robot (see Fig. 12). Second, reconfiguring the pins to allow 90-degree rotation of the connector would eliminate the need for the rotating table. Third, efforts should be made to increase clearance near the pin. This could be achieved by removing the screw beside the pins, which prevents the gripper from opening properly (see Fig. 13). Fourth, the color of the connector should be different from the background objects to increase the accuracy of pose estimation (see left side of Fig. 14). The application of texture on the connector would also facilitate image processing (see the right side of Fig. 14).

TABLE I
32 TRIALS OF CONNECTOR GRASPING AND INSERTION

Number	X(pixels)	Y(pixels)	θ (°)	Overall
1.	79.86	82.77	7.8533	☑
2.	98.1575	73.84	9.46	☑
3.	82.4719	80.6225	-9.782	☑
4.	103.469	85.5117	-14.534	X
5.	103.031	89.042	-11.689	☑
6.	90.323	91.418	17.818	☑
7.	86.203	82.397	9.782	X
8.	90.257	86.244	-19.025	☑
9.	91.227	85.111	-3.691	☑
10.	89.874	88.187	12.094	☑
11.	86.672	79.381	14.036	☑
12.	91.877	87.584	23.962	☑
13.	91.064	90.441	3.945	X
14.	85.561	94.070	5.710	☑
15.	85.904	86.815	-16.504	☑
16.	81.391	82.961	0	☑
17.	90.570	89.961	14.036	☑
18.	96.636	72.570	0	☑
19.	103.214	82.367	24.7751	☑
20.	88.964	85.785	-3.945	☑
21.	88.594	100.823	0	☑
22.	96.595	80.680	-11.309	☑
23.	84.003	85.029	-21.447	☑
24.	72.665	104.918	0	☑
25.	77.328	88.233	3.945	☑
26.	76.031	95.981	3.814	☑
27.	100.364	77.550	3.814	☑
28.	85.828	86.871	0	☑
29.	68.718	89.927	0	☑
30.	69.9068	83.387	-11.309	☑
31.	92.206	77.114	-3.814	☑
32.	80.924	95.723	24.775	☑

TABLE II
ELAPSED TIME FOR EACH OPERATION

ACTION	TIMELINE (Sec)	TIME (Sec)
Home - Grasping	(0 - 07)	7
Grasping	(07 - 11)	4
Grasping- complete 180-degree flip	(11 - 22)	11
180-degree flip- Before rotation	(22 - 28)	6
Rotating table rotation	(28 – 33)	5
Completion of rotation - Insertion	(33 - 39)	6
Push operation	(39 – 44)	5
Return home (robot)	(44 – 52)	8
Rotation table rotating back	(52 - 57)	5

The factory environment could be improved by reducing the camera to object distance and switching to a lens with a wider field of view (FOV). It would also be possible to use a camera with lower resolution in order to reduce costs. The performance of the system should not be compromised as long as the minimum resolution is met; i.e., mm (x)/pixel(x) = 0.3 to 0.36 (along X axis) mm and (y)/pixel(y) = 0.2 to 0.28 (along Y axis).

Fig. 11. Limit on rotation angle of connector around Z-axis: 25 degrees (up) and -25 degrees (down)

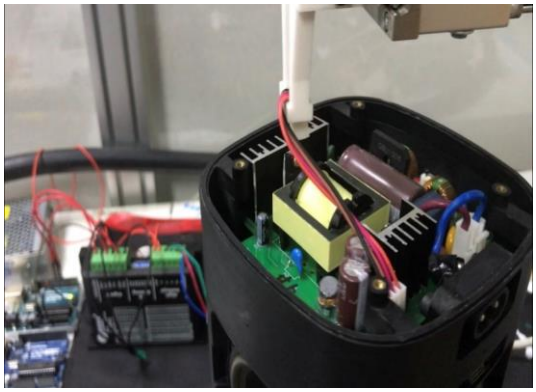


Fig. 12. Short length of connector wires hindering insertion

V. CONCLUSIONS

This paper outlines a novel approach to the grasping and insertion of connectors by a robot. We employed two 2D cameras with various image-processing techniques (binary threshold, closing and opening, and contouring) to enable the detection of a connector elevated in the air by wires. The proposed system is far less expensive than systems based on structured-light 3D cameras. In experiments involving 32 trials, the proposed scheme achieved a success rate of 90%. A number of suggestions for improvement are provided to enhance the accuracy of the system in future research.

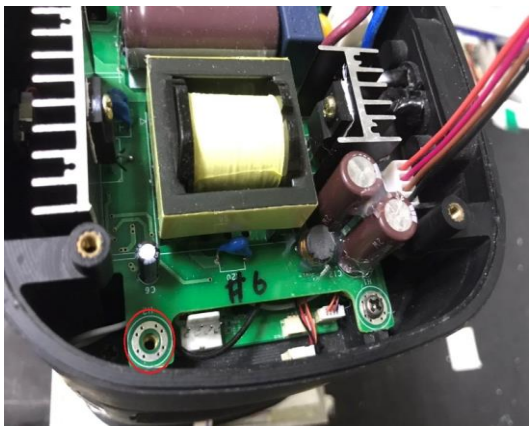


Fig. 13. Screw placement hindering insertion

Fig. 14. Color of connector wire hindering insertion

REFERENCES

[1] S.-I. Choi and S.-Y. Park, "A new 2-point absolute pose estimation algorithm under planar motion," Adv. Robot., vol. 29, no. 15, pp. 1005-1013, 2015.

[2] C. Choi and H. I. Christensen, "Real-time 3D model-based tracking using edge and keypoint features for robotic manipulation," in Proc. IEEE Int'l Conf. Robotics Automation (ICRA), 2010, pp. 4048-4055.

[3] K. T. Song and C. H. Chang, "Object pose estimation for grasping based on robust center point detection," in Control Conference (ASCC), 2011 8th Asian, 2011, pp. 305-310.

[4] Gocator. "3D snapshot sensors." Gocator.com.

https://lmi3d.com/products/gocator/g3/snapshot-sensors (accessed Sep. 9, 2020).

[5] Solomon. "Robot applications." Solomon.com.tw

https://www.solomon-3d.com (accessed Sep. 9, 2020).

[6] L. Yang, E. Li, T. Long, J. Fan, and Z. Liang, "A novel 3-D path extraction method for arc welding robot based on stereo structured light sensor," IEEE Sensors J., vol. 19, no. 2, pp. 763-773, Jan. 2018.

[7] Y. Chen, G. Sun, H. Lin, and S. Chen. "Random bin picking with multiview image acquisition and cad-based pose estimation," in 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 2218-2223, Oct 2018.

[8] D. Li, N. Liu, Y. Guo, X. Wang, and J. Xu, "3D object recognition and pose estimation for random bin-picking using Partition Viewpoint Feature Histograms," Pattern Recognition Letters, vol. 128, pp. 148-154, 2019.

[9] S. Diao, X. Chen, J. Luo, "Development and Experimental Evaluation of a 3D Vision System for Grinding Robot," Sensors 2018, vol. 18, no. 9, 3078.

[10] H. I. Lin and C. M. Nguyen, "Boosting Minority Class Prediction on Imbalanced Point Cloud Data," Applied Sciences, vol. 10, no. 3, 973, 2020.

[11] R. B. Rusu, N. Blodow, and M. Beetz, " Fast point feature histograms (FPFH) for 3D registration," in IEEE international conference on robotics and automation, 2009, pp. 3212-3217.

[12] Z. Wang, J. Fan, F. Jing, Z. Liu, and M. Tan, "A pose estimation system based on deep neural network and ICP registration for robotic spray painting application," in int J. Adv. Manuf. Technol. vol 104, no 1, pp. 285–299, 2019.

[13] S. Ying, J. Peng, S. Du, and H. Qiao, "A scale stretch method based on ICP for 3D data registration," IEEE Trans. Autom. Sci. Eng. vol. 6, no 3, pp. 559-565, 2009.

[14] J. Chen, X. Wu, M. Y. Wang, and X. Li, "3D shape modeling using a self-developed hand-held 3D laser scanner and an efficient HT-ICP point cloud registration algorithm," Opt. Laser Technol. vol. 45, pp. 414-423, 2013.

[15] J. M. Phillips, R. Liu, and C. Tomasi, "Outlier robust ICP for minimizing fractional RMSD," in the sixth international conference on 3-D Digital Imaging and Modeling, pp. 427-434, 2007.